TAG CLOUDS FOR SITUATED INTERACTION AND PLACE PROFILING

Rui José, Bruno Silva
Information Systems Department, University of Minho, Guimarães, Portugal
{rui, bruno.silva}@dsi.uminho.pt

Fernando Reinaldo Ribeiro
Informatics Department, Polytechnic Institute of Castelo Branco, Castelo Branco, Portugal
fribeiro@ipcb.pt

Keywords: Tag clouds, Situated interaction, Public displays, Bluetooth, Places, Pervasive computing.

Abstract: Tag clouds have become very popular as visual representations of the main topics in document sets or as navigation tools that can provide quick access to resources related with specific topics. However, their ability to represent the information environment associated with any meaningful reality in a way that is collectively visible, actionable and easily understood may also be very relevant, even when the reality being represented is no longer a set of documents or resources, but a stream of interactions occurring within a particular ubiquitous computing environment. In this paper, we explore the use of tag clouds within the context of situated displays and services. We hypothesise that such tag clouds may have a role as dynamic representations of place and also as interaction controls, supporting the same comprehension and navigation functions of classical tag clouds. We describe two case studies in which this concept of situated tag cloud has been experimented in real-world settings. The case studies demonstrate two different applications of the tag cloud concept as the basis for place description and situated interaction. The results obtained from the case studies suggest that situated tag clouds can indeed provide valuable representations of place and situations and can also support simple interaction models, allowing people to reason about the system behaviour and how it is being influenced by new interactions.

1 INTRODUCTION

A tag cloud is a visualization of a weighted list of words in which those weights are associated with visualization attributes, typically size and colour. Tag clouds have become very popular as a visualization mechanism for the main topics associated with a web site or set of documents. The tags are obtained from the words in the text and the size corresponds to the relative frequency of each word. They offer a visual representation of the main topics in that text, providing a simple and yet powerful perspective of the respective content. Tag clouds are also extensively used in crowdsourcing systems where people can freely tag photos, web sites, videos or other resources. In this case, the generation of the tag cloud becomes a collective process of knowledge creation and the tag cloud may become a navigation tool, providing quick access to resources related with specific topics.

What is very powerful about the concept of tag clouds is their ability to represent the information environment associated with any meaningful reality in a way that is collectively visible, actionable and easily understood. They may facilitate comprehension by offering an alternative and instant illustration of the main topics associated with a particular entity, such as a document, a web site, or a person's interests. They may also facilitate navigation by providing an alternative to more conventional navigation patterns, such as those based on menus. In other words, they provide a social affordance in the sense that they convey social interaction of fellow users and potentials for social interaction (Bielenberg and Zacher, 2005). These are valuable features that may still be useful, even if the reality being represented is no longer a set of documents or resources, but a stream of interactions occurring within a particular ubiquitous computing environment. As stated by Joe Lamantia (Lamantia, 2006) , "*tag clouds are revolutionary in their ability to translate the concepts associated with nearly anything you can think of into a collectively visible and actionable information environment*".

Our aim is to explore the use of tag clouds within the context of situated displays and services. In these environments, people may interact with physically situated services, such as public displays, Bluetooth hotspots or location-based services. This type of interaction has the potential to generate words, which in turn may be used to dynamically generate an evolving tag cloud that somehow represents those situated interactions.

We hypothesise that such tag clouds may have a role as dynamic representations of place and also as interaction controls, supporting the same comprehension and navigation functions of classical tag clouds. They may support comprehension by offering a representation of the topics associated with a place. They may also support navigation by offering an aggregate interaction mechanism, whereby keywords in the tag cloud provide an immediately available and dynamically evolving list of interaction suggestions. The tag cloud concept may thus be used as a unifying concept for the association between keyword inputs and dynamic behaviours. It can provide a visualization that may be relevant in itself as a particular characterization of place and at same time work as a driver for the display behaviour by enabling the selection or generation of situationally relevant content.

In this paper, we describe a system for generating place-based tag clouds and two case studies in which this concept of place-based tag cloud has been experimented in real-world settings. The results suggest that situated tag clouds can indeed play a role as dynamic representations of place and also as interaction controls, although there are some open issues with the interaction models.

2 RELATED WORK

Many researchers have reported alternative applications of the tag cloud concept to represent realities outside their original context, including personal profiles, social dynamics in meetings or trends in political discourse.

Steinbock et al. studied the use of wearable tag clouds in face-to-face interaction (Steinbock et al., 2007). Within the context of an academic meeting, participants were given a large badge with a tag cloud of the most common words in their published documents. This was expected to represent a synopsis of the respective interests and facilitate interaction between participants. McNaught and Lam (McNaught and Lam, 2010) explored tag clouds to analyse the spoken and written responses of informants in focus groups transcripts. The tag clouds facilitate the study of the social dynamics in those groups, but this analysis is only conducted at a later stage and provides no feedback on the social interaction as it unfolds. Viégas and Wattenberg report on some of the uses of IBM web site *Many Eyes*, where people upload and visualize data in a variety of ways (Viegas and Wattenberg, 2006). One of those visualizations enables people to represent their profile through a tag cloud generated from the set of blogs that person normally reads. Tagline Generator (Mehta, 2006) supports the generation of chronological tag clouds from time-based text data sources. This work is an interesting example of a system that deals with the time dimension. In these tag clouds, the colour in the words is associated with usage variations. Words whose usage is increasing will be brightened, while words whose usage is decreasing will be fading away. The use of tag clouds for content recommendation has been described by Pessemier et.al (Pessemier et al., 2009). Tag clouds are generated from user ratings to create a form of personal profile. These tag clouds are then used to recommend movies to that person. In our work, we also suggest the use of tag cloud for recommendation purposes, but in our case this corresponds to a place or situation profile, and not to the profile of a single individual.

This related work demonstrates the variety of applications for the tag cloud concept. However, the use of tag clouds as a situated representation of place remains to be explored.

3 SITUATED TAG CLOUDS

A situated tag cloud is a tag cloud generated from the keywords extracted from implicit and explicit interactions observed in the context of a particular situated system. The most distinguishing property in the case of situated tag clouds is the way the input words are obtained. Instead of counting the words within a particular set of documents, we have a continuous stream of words being generated by various types of situated interactions, such as Bluetooth names, Obex exchanges or SMS/MMS messages. Additionally, we have a specific physical and social setting within the context of which we need to interpret those words. As part of this research, we have developed a system for creating tag clouds based on the digital footprints generated from presence and interaction events associated with a public display. The tag cloud generation mechanism is part of a larger system, called instant places, in which people can use their Bluetooth device name for managing their presence and

activating interactive features. The details of instant places are outside the scope of this paper, but the reader is referred to (José et al., 2008) for a more detailed account of the system characteristics.

3.1 Situated Identities

The instant places system recognizes the presence of Bluetooth devices and generates persistent identities that may evolve with the history of presence and interaction in that particular place. The system also recognizes parts of the Bluetooth device name as explicit instructions, allowing people to use their Bluetooth name to trigger specific behaviours on the situated display, such as the presentation of specific content. The resulting data model, composed by the observed identities and their actions, is constantly being shaped by new interaction and presence events and is used as the main driver in the behaviour of the situated display. Identities that are currently present are the major source for the generation of tags. The presence of tags represents an additional dimension that is not normally included in traditional tag clouds, but may be key to introduce the situatedness of a situated tag cloud. Even tag clouds with a large time span may thus favour tags that are currently present, albeit less popular, instead of popular tags that no one is currently generating.

3.2 Token Data Model

The list of currently present identities constitutes a dynamic keyword source that can be sampled periodically to generate an evolving collection of place keywords. The token data model is the result of this sampling process in which keyword snapshots are periodically being generated.

A keyword snapshot is made of a list of keywords and their corresponding presence intensity. The generation of this list can be tailored using a presence level map that associates specific command types with various weights in their contribution to the intensity of presence. Because situated tag clouds are generated as part of specific interaction events and also because they can be created to support specific forms of adaptation, they can be optimised for different purposes and to represent multiple dimensions of place. For example, a music tag cloud may be created with words and interactions related with music and be used to support music recommendation for the place.

3.3 Tag Cloud Specification

The Tag cloud module supports the creation of tag clouds that will be associated with a particular place. Each tag in a tag cloud has four key parameters: A name corresponding to the tag itself; a popularity corresponding to the accumulated presence of that tag as observed in the keyword snapshots within a particular time scope; a presence level indicating the current presence level for that tag; and a rank that arranges the most popular keywords, the ones that are actually making it into the tag cloud, in a number of categories according to their popularity.

The creation of a new tag cloud is specified by setting the key parameters that will determine the tag cloud behaviour. These parameters may be arranged in four major blocks: General, time dynamics, place-making and visualization.

3.3.1 General Parameters

General parameters essentially serve to identify and describe the tag cloud. They include a name that is a unique identifier within the respective place. There is also a title that describes the overall concept of the tag cloud in terms of its source and dynamic properties. The main purpose is to enable people to interpret the tag and reason about it, and thus it should be presented when the tag cloud is visualised. Additional parameters include the number of tags to be included, the rank Algorithm that defines the model used for distributing the N most popular tags among the ranks, the number of rank levels, and the minimum word length.

3.3.2 Time Dynamics

Tag clouds are normally generated from a reasonably static resource set, meaning that their frequent update is not a major concern. With situated tag clouds, time gains an increased importance because the source for words is a continuous sequence of presence and interaction events. Time is therefore a key part of the underlying data model for situated tag clouds. After some time, data becomes less relevant and eventually it should be discarded. A particular memory span defines for how long the tag cloud will retain the sighted tags before the information is discarded.

Time dynamics parameters specify the tag cloud behaviour in relation with time and the deprecation of the keyword observations. Managing time involves defining a strategy for two main issues: the minimum time scale for keyword aggregation and the deprecation policy for discarding or reducing the weight of older tag sightings. The combination of these two policies will determine whether the system is very reactive to the presence of new tags, or is a more stable representation of place. They can be set

through the following parameters: The *Period Type* determines the time aggregation level from which the tag cloud is calculated. The *Number of Periods* defines the number of time periods of type *Period Type*. Information deprecation is controlled by setting the *Period Decay* parameter that progressively reduces the weight of older periods in the final aggregate information that is going to constitute the tag cloud.

3.3.3 Place-making

The situated tag cloud must integrate specifications that align its evolution with the place-making goals of whoever installed the display and should support an implicit negotiation between the place owner perspective and the input from multiple visitors.

Place-making parameters allow a place owner to provide additional characterization and specify adaptation boundaries for a tag cloud. Even though these specifications are not meant to determine the final outcome of the tag cloud, they provide a way for aligning the display behaviour with the general expectations of appropriateness of the display owner and its place-making objectives. The main part of these specifications is a set of keyword lists that enable some control of the tags in the tag cloud, including a blocked words list or a list of seed keywords that serve to initialise the tag cloud and maintain a number of place keywords when there are not enough keywords being generated. Seed tags are defined with a minimum popularity value that determines how visible they remain when other tags begin to emerge. A *Seeds Only* parameter can be used to determine that a particular tag cloud will only accept seed words. This works as a white list that restrains the accepted words to those on the list. This closed tag cloud model may be useful for thematic tag clouds, e.g. a tag cloud with sport teams, emoticon symbols, or music styles. A tag cloud based on a white list promotes aggregation around particular tags. Finally, there is also a list of contextual keywords that can be used to provide additional context to the words in the tag cloud. For example, Sports and Football could be added to a tag cloud representing football teams. This is particularly important if the tag cloud is to be used for selecting web content.

3.3.4 Visualization

Visualization is an integral part of any tag cloud. In its essence, the way a place-based tag cloud may be presented is basically the same as for any other tag cloud. In the basic format, tags are primarily sorted alphabetically with the most important items being shown with larger font sizes. The main difference is that current presence is highlighted through the use of colour. Another difference is the possibility, in the case of closed tag clouds, of using images instead of words, even if each image is directly associated with a specific word.

4 EVALUATION

We have explored this concept of situated tag cloud in two separate case-studies, which we will now briefly describe.

4.1 Scheduling for Public Displays

In our first case study, we have used a generic tag cloud to represent a dynamic and evolving view of a place that could then be used as the key input to a recommender system that would select information feeds for presentation in a public display. The full details of this case-study can be found in (Ribeiro and José, 2010), but for the convenience of the reader, we briefly summarise the key points of that case study.

A context-aware scheduler considered both the weight and current presence level of the represented tags to select which content it would show next. The tag cloud was specified by the place owner, but place visitors could publish their own interests through tag commands in their Bluetooth device name. The combination of multiple contributions from place visitors was then used for content recommendation on the public display.

The evaluation of this system was based on a 3 weeks experiment in which the tag cloud was seeded with 20 words representing topics related with Informatics and Engineering as well as location related keywords associated with the town and region. During these 3 weeks we collected usage logs and conducted a total of 15 structured interviews with people who had previously tried to use the system. The results obtained with this case study suggest that this is a viable approach to the problem of selecting relevant content for a dynamic view of place. In particular, the visual nature of the tag cloud seemed to facilitate the interpretation of the system behaviour in a way that influences positively the user perception, even when the selections are not perfect. The results have also shown that place visitors recognize the sensitivity of the system to their demands and that a place tag cloud can provide an important element for the interpretation of place and for the combination of the interests expressed by the place owner and the mul-

tiple place visitors.

4.2 Group Expression at School

In our second case-study, we deployed a public display in a school in which tag clouds were used to represent the popularity of local communities, such as students' classes and teachers' unit departments. We had two separate tag clouds: one for students, with 56 seed tags, and the other for teachers, with 12 seed tags. The *seeds only* property was set to true, so that only the seed tags appeared in the display. The popularity was calculated from the sightings of the last two weeks, with the current week being counted 100%, and the previous one only 50%. The display was divided in four panels, as shown in figure 1.

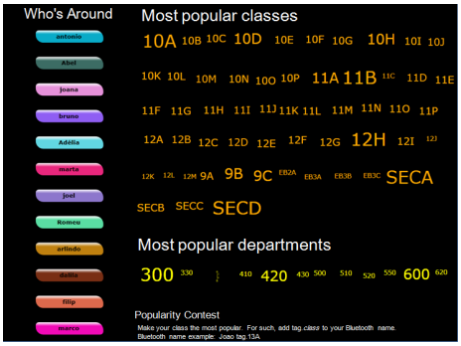

Figure 1: Group presence representation.

The left panel displays the Bluetooth device names that are currently detected in the place. The top right panel represents the students' classes tag cloud and the low right panel represents the professors' unit departments tag cloud. At the bottom, instructions for using the tag command were provided.

Although this experience was part of a broader study lasting twenty-one weeks, this specific experience was run for a period of four weeks. During those four weeks, 313 different devices were detected and 25 of them used the tag command, which represents 7.98% of the total device detection. In this period, there were 3226 sessions, in which 259 were produced by the devices that used the tag command, which represents 8.02%. In terms of individual tag cloud interaction, 23 devices in 44 session interacted with the students' tag cloud and 2 devices in 215 sessions interacted with the professors' tag cloud.

This particular type of situated tag cloud seems to have generated an interaction pattern that was radically different from the previous case study. After an initial period with no interactions, the first interactions appeared and seemed to have sparred a significant set of additional interactions. What we have observed from these numbers and also from the interviews was that this was a case in which early adopters were hard to get. Only when someone started pushing their group name, would others, from the same or other groups, would follow, almost as a reaction.

4.3 Analysis

The main objective of this work was to explore how we can leverage on the tag cloud concept to support situated interaction. One way to answer this question is to consider to what extent our notion of tag cloud is in fact similar to the traditional notion of tag cloud, and to what extent existing tools for the generation of situated keywords may still be used to create situated tag clouds.

Clearly the underlying data model for generating and storing words is very different, given the substantial differences in the words generation processes. However, independently of the data model being used for managing the keywords obtained from situated interaction, it would be trivial to generate an equivalent text based containing those same words, or at least the most popular ones, with each word being represented as many times as any internal frequency equivalent. This would allow existing tools to take that document as their input and generate the tag cloud, which means that at least from the perspective of generating the visualizations there is some potential for leveraging on existing tools. There are however, some specificities that do not map well into this model. In particular, presence information represents an additional information attribute that cannot be captured by tag frequency alone and that is not considered by most tag cloud tools. While many tag cloud generators support the use of colours, for example, as an additional visual element, this is only used for aesthetic purposes and is not linked to any variable.

Regarding the case-studies, they demonstrate various ways in which tag clouds may be used to represent place and support situated interaction. In the first case-study the tag cloud is continuously reflecting the social setting around the display, being sensitive to immediate indications of interest and providing a balanced combination between the content suggestions expressed by multiple place visitors and those expressed by the place owner. In the second case-study it represents the flow of groups around the display setting.

5 CONCLUSIONS

Tag clouds as proposed in this work exhibit very interesting capabilities for supporting a balanced combination of information filtering and information retrieval. They support information filtering because even if no one is using the system, the tag cloud is already there and capable to serve as a content generator. Moreover, the tag cloud specification defines an adaptation scope that limits the extent of the content that can be displayed. They support post-filtering because once the display is in place, people passing-by will sort out the content deemed more interesting to them. Moreover, the tag cloud provides an interesting representation of the interests of a crowd. It avoids merely determining averages that are not representative. It can also deal with the tension between place adaptation, as something that can be learned over time, and situatedness, as the ability to react quickly to the social dynamics around the display. Another very positive point is the way in which the tag clouds can be visualised and enhance the perception that people may have of the adaptation processes going on.

We also found that peoples' expectations may not be aligned with the interaction concepts upon which the model is based. While the place-based tag cloud is essentially designed as crowd interaction mechanism, and, moreover framed within the concept of place, people often expect the system to exhibit an immediate reaction to their specific interaction. Moreover, the context and the semantics of tagging in this context are ambiguous. When someone advertises a tag that is collected by the system at a particular place, what are they tagging?: the place, themselves or that particular situation? Perhaps people do not even think of themselves as tagging, but rather as interacting with a system that accepts words as input. In either case, peoples' perception about these issues and the tagging patterns that may emerge will necessarily have a major impact on the viability of this approach.

5.1 Future Work

Further research is needed to evaluate across multimultiple settings the ideal values for some of the system parameters. For example the decay of user-suggested tags affects responsiveness and also the balance between pre-defined and emerging notions of place, while the size of non-repetition queues affects the balance between content quality and diversity. Results suggest that this may be a valuable step towards the emergence of dynamic place profiles that match the social expectations and practices of their evolving social settings. Following on this idea, we also intend to explore how the similarity between places can be inferred from the similarity between the respective tag clouds.

ACKNOWLEDGEMENTS

Fernando Ribeiro was supported by a Portuguese Foundation for Science and Technology scholarship (SFRH/BD/31292/2006).

The research leading to these results has received funding from FCT under the Carnegie Mellon - Portugal agreement: Wesp (Web Security and Privacy (Grant CMU-PT/SE/028/2008).

REFERENCES

Bielenberg, K. and Zacher, M. 2005, 'Groups in Social Software:Utilizing Tagging to Integrate Individual Contexts for Social Navigation', Master's thesis thesis.

José, R., Otero, N., Izadi, S. and Harper, R., 2008. Instant Places: Using Bluetooth for Situated Interaction in Public Displays. *IEEE Pervasive Computing*, vol. 7, no. 4, Oct.-Dec., pp. 52-7.

Lamantia, J., 2006. *Tag Clouds Evolve: Understanding Tag Clouds*. Joe Lamantia.com Retrieved from: http://www.joelamantia.com/ideas/tag-clouds-evolve-understanding-tag-clouds, viewed September 22.

McNaught, C. and Lam, P., 2010. Using Wordle as a supplementary research tool. The Qualitative Report. vol. 15(3), no. 3, pp. 630-43, viewed September 2010, <http://www.nova.edu/ssss/QR/QR15-3/mcnaught.pdf>.

Mehta, C., 2006. *Tagline Generator*. Retrieved from: http://chir.ag/projects/tagline, viewed Sept. 2010.

Pessemier, T. D., Deryckere, T. and Martens, L., 2009. Context aware recommendations for user-generated content on a social network site. *7th European interactive television conf.* Leuven, pp. 133-6.

Ribeiro, F. and José, R. 2010, 'Autonomous and Context-Aware Scheduling for Public Displays Using Place-Based Tag Clouds', in Springer-Verlag (ed.), *Ambient Intelligence and Future Trends-International Symp. on Ambient Intelligence*, vol. 72, pp. 131-8.

Steinbock, D., Pea, R. and Reeves, B., 2007. Wearable tag clouds: visualizations to facilitate new collaborations *8th Int. Conf. on Computer supported collaborative learning.* New Brunswick, New Jersey, pp. 672-4.

Viegas, F. and Wattenberg, M., 2006. Communication-Minded Visualization: A Call to Action. *IBM Systems Journal*, vol. 45, no. 4, pp. 801 - 12.